Please amend the remaining claims as follows:

1. (Once amended) A vector construct comprising:

(a) a first transcriptional regulatory sequence operably linked to a first unpaired splice donor sequence; and

(b) a second transcriptional regulatory sequence operably linked to a second unpaired splice donor [sequence; and

(c) a linearization site.] sequence.

2. (Once amended) The vector construct of claim 1, wherein said [linearization site is located between said first unpaired splice donor site and] first transcriptional regulatory sequence is in the same orientation as said second transcriptional regulatory sequence.

3. (Once amended) The vector construct of claim 1, wherein [when said vector integrates into the genome of a host cell,] said first transcriptional regulatory sequence is in an inverted orientation relative to the orientation of said second transcriptional regulatory sequence.

4. (Once amended) The vector of claim [1,] 2 or claim 3, wherein said vector has been rendered linear [by cleavage at said linearization site].

6. (Once amended) A vector construct comprising, in sequential order:

(a) a transcriptional regulatory sequence;

(b) [a vector-encoded] an exon comprising a rare cutting restriction site;

(c) an unpaired splice-donor site; and

(d) a linearization site.

7. (Once amended) [A] The vector construct of claim 6, further comprising[, in sequential order:

(a) a transcriptional regulatory sequence;

(b) a vector-encoded exon comprising a first rare cutting restriction site;

(c) an unpaired splice-donor site;

(d)] a second rare cutting restriction site located between said unpaired splice donor site and said linearization site.[; and

(e) a linearization site.]

11. (Once amended) The vector construct of [any one of claims 1, 8, or 10,] claim 1 or claim 10, wherein said first transcriptional regulatory sequence or said second transcriptional regulatory sequence is a promoter.

13. (Once amended) The vector construct of any one of claims 5-7 [or 9], wherein said transcriptional regulatory sequence is a promoter.

15. (Once amended) The vector construct of [any one of claims 8-10] claim 10, wherein said selectable marker is selected from the group consisting of a neomycin gene, a hypoxanthine phosphribosyl transferase gene, a puromycin gene, a dihydrooratase gene, a glutamine synthetase gene, a histidine D gene, a carbamyl phosphate synthase gene, a

dihydrofolate reductase gene, a multidrug resistance 1 gene, an aspartate transcarbamylase gene, a xanthine-guanine phosphoribosyl transferase gene, an adenosine deaminase gene, and a thymidine kinase gene.

20. (Once amended) A eukaryotic host cell comprising the vector construct of any one of claims [1, 5-10, or 16.] 1, 5, 6, 7 and 10.

34. (Once amended) The primer molecule of claim 33, wherein said hapten molecules are selected from the group consisting of biotin, digoxigenin, an antibody, an enzyme, lipopolysaccharide, apotransferrin, ferrotransferrin, insulin, a [cytokine] cytokine, an extracellular matrix protein, an integrin, ankyrin, C3bi, fibrinogen, spectrin, a cytokine receptor, an insulin receptor, a transferrin receptor, polymyxin B, endotoxin-neutralizing protein (ENP), an enzyme-specific substrate, protein A, protein G, a cell-surface Fc receptor, an antibody-specific antigen, an antibody-specific peptide, avidin, and streptavidin.

36. (Once amended) A method for first strand cDNA synthesis comprising:

(a) annealing [a] the primer of claim 30 to an RNA template molecule to form an primer-RNA complex; and

(b) treating said primer-RNA complex with reverse transcriptase and one or more deoxynucleoside molecules under conditions favoring the reverse transcription of said primer-RNA complex to synthesize a first strand cDNA.

Please insert the following new claims:

--58. A vector construct comprising a transcriptional regulatory sequence operably linked to an unpaired splice donor sequence and one or more amplifiable markers, wherein said vector construct does not comprise a homologous targeting sequence.

59. A vector construct comprising a transcriptional regulatory sequence, an amplifiable marker, and a viral origin of replication.

60. A vector construct comprising a selectable marker, a transcriptional regulatory sequence operably linked to a translational start codon, a secretion signal sequence, an epitope tag, and an unpaired splice donor site.

61. A vector construct comprising a transcriptional regulatory sequence operably linked to a translational start codon, a secretion signal sequence, an epitope tag, a sequence-specific protease site, and an unpaired spice donor site.

62. A vector comprising:

(a) a transcriptional regulatory sequence operably linked to a translation start codon,

(b) a nucleic acid sequence encoding an amino acid sequence of four or more amino acids, wherein said amino acid sequence alone is insufficient to constitute signal peptide activity, but is sufficient to constitute signal

peptide activity when said nucleic acid sequence is combined with or is upstream of an exon of an endogenous gene, and

(c) an unpaired splice donor site.

63. The vector construct of any one of claims 60-62, wherein said construct further comprises one or more amplifiable markers.

64. The vector construct of any of claims 58 and 60-62, wherein said transcriptional regulatory sequence is a promoter.

65. The vector construct of claim 64, wherein said promoter is a viral promoter.

66. The vector construct of claim 65, wherein said viral promoter is a cytomegalovirus immediate early gene promoter.

67. The vector construct of claim 65, wherein said promoter is a non-viral promoter.

68. The vector construct of claim 65, wherein said promoter is an inducible promoter.

~~69. A cell containing the vector construct of any one of claims 58-62.~~

70. A cell containing the vector construct of claim 63.

71. The cell of claim 69, wherein said vector construct has integrated into the cellular genome.

72. The cell of claim 70, wherein said vector construct has integrated into the cellular genome.

73. The cell of claim 71 or 72, wherein an endogenous gene is over-expressed in said cell by upregulation of the gene by said transcriptional regulatory sequence on said vector construct.

74. The cell of claim 69, wherein said cell is an isolated cell.

75. The cell of claim 70, wherein said cell is an isolated cell.

76. A method for making a host cell, comprising introducing the construct of any one of claims 58-62 into a cell.

77. A method for producing an expression product of an endogenous cellular gene or portion thereof comprising:

(a) introducing the construct of any one of claims 58-62 into a genome-containing cell;

(b) integrating said construct into the genome of said cell by non-homologous recombination; and

(c) over-expressing said endogenous gene in said cell.

78. The method of claim 77, wherein said over-expression is accomplished *in vitro*.

79. The method of claim 77, wherein said over-expression is accomplished *in vivo*.

80. The method of claim 77, further comprising isolating said expression product from said cell.

81. A cell library comprising a collection of cells transformed with the construct of any one of claims 58-62, wherein said construct is integrated into the genomes of said cells by non-homologous recombination.

82. A method of obtaining a gene product from a library of cells comprising screening the library of claim 81 for expression of said gene product, selecting from said library a cell that over-expresses said gene product, and obtaining said gene product from said selected cell.

83. A method for producing an expression product of an endogenous cellular gene comprising:

(a) introducing a vector comprising a transcriptional regulatory sequence operably linked to a secretion signal sequence and an unpaired splice donor sequence into a cell;

(b) integrating said vector into the genome of said cell by non-homologous recombination;

(c) over-expressing an endogenous gene or a portion thereof in said cell by upregulation of said gene by said transcriptional regulatory sequence;

(d) screening said cell for over-expression of said endogenous gene or portion thereof; and

(e) culturing said cell under conditions favoring the production of the expression product of said endogenous gene or portion thereof by said cell.

84. The method of claim 83, further comprising isolating said expression product.

85. A method for over-expressing an endogenous gene in a cell *in vivo*, comprising:

(a) introducing a vector comprising a transcriptional regulatory sequence into a cell;

(b) integrating said vector into the genome of said cell by non-homologous recombination;

(c) over-expressing an endogenous gene or a portion thereof in said cell by upregulation of said gene by said transcriptional regulatory sequence;

(d) screening said cell for over-expression of said endogenous gene; and

(e) introducing said isolated and cloned cell into an animal under conditions favoring the overexpression of said endogenous gene by said cell *in vivo*.

Sub D5

86. A method for producing an expression product of an endogenous cellular gene *in vivo*, comprising

(a) introducing a vector comprising a transcriptional regulatory sequence operably linked to an unpaired splice donor sequence into a cell;

(b) integrating said vector integrate into the genome of said cell by non-homologous recombination;

(c) over-expressing an endogenous gene or a portion thereof in said cell by upregulation of said gene by said transcriptional regulatory sequence;

(d) screening said cell for over-expression of said endogenous gene; and

(e) introducing said isolated and cloned cell into an animal under conditions favoring the overexpression of said endogenous gene by said cell *in vivo*.

Sub D5

87. A method for producing an expression product of an endogenous cellular gene, comprising:

(a) introducing a vector comprising a transcriptional regulatory sequence and one or more amplifiable markers into a cell;

(b) integrating said vector into the genome of said cell by non-homologous recombination;

(c) over-expressing an endogenous gene or a portion thereof in said cell by upregulation of said gene by said transcriptional regulatory sequence;

(d) screening said cell for over-expression of said endogenous gene;

(e) culturing said cell under conditions in which said vector and said endogenous gene are amplified in said cell; and

(f) culturing said cell under conditions favoring the production of the expression product of said endogenous gene by said cell.

88. The method of claim 87, further comprising isolating said expression product.

89. The method of claim 87, wherein said vector further comprises a splice donor site operably linked to said transcriptional regulatory sequence.

90. The method of claim 87 or claim 89, wherein said endogenous gene or portion thereof encodes a protein selected from the group of proteins consisting of erythropoietin, insulin, growth hormone, glucocerebrosidase, tissue plasminogen activator, granulocyte-colony stimulating factor (G-CSF), granulocyte/macrophage colony stimulating factor (GM-CSF), macrophage colony-stimulating factor (M-CSF), interferon α, interferon β, interferon γ, interleukin-2, interleukin-3, interleukin-4, interleukin-6, interleukin-8, interleukin-10, interleukin-11, interleukin-12, interleukin-13, interleukin-14, TGF-β, blood clotting factor V, blood clotting factor VII, blood clotting factor VIII, blood clotting factor IX, blood clotting factor X, TSH-β, bone growth factor-2, bone growth factor-7, tumor necrosis factor, alpha-1 antitrypsin, anti-thrombin III, leukemia inhibitory factor, glucagon, Protein C, protein kinase C, stem cell factor, follicle stimulating hormone β, urokinase, a nerve growth factor, an insulin-like growth factor, insulinotropin, parathyroid hormone, lactoferrin, a complement inhibitor, platelet derived growth factor, keratinocyte growth factor, hepatocyte growth factor, endothelial cell growth factor, neurotropin-3, thrombopoietin, chorionic gonadotropin, thrombomodulin, alpha glucosidase, epidermal growth factor, fibroblast growth factor, a cell surface receptor, a

transmembrane ion channel, a cholesterol receptor, a receptor for a lipoprotein, an integrin, a cytoskeletal anchoring protein, an immunoglobulin receptor, and a CD antigen.

91. The method of claim 87 or claim 89, wherein said endogenous gene or portion thereof encodes an erythropoietin protein.

92. The method of claim 87 or claim 89, wherein said endogenous gene or portion thereof encodes a growth hormone protein.

93. The method of claim 87 or claim 89, wherein said endogenous gene or portion thereof encodes a G-CSF protein.

94. A gene expression product produced by the method of claim 87 or claim 89, wherein said gene expression product is a protein selected from the group of proteins consisting of erythropoietin, insulin, growth hormone, glucocerebrosidase, tissue plasminogen activator, granulocyte-colony stimulating factor (G-CSF), granulocyte/macrophage colony stimulating factor (GM-CSF), macrophage colony-stimulating factor (M-CSF), interferon α, interferon β, interferon γ, interleukin-2, interleukin-3, interleukin-4, interleukin-6, interleukin-8, interleukin-10, interleukin-11, interleukin-12, interleukin-13, interleukin-14, TGF-β, blood clotting factor V, blood clotting factor VII, blood clotting factor VIII, blood clotting factor IX, blood clotting factor X, TSH-β, bone growth factor-2, bone growth factor-7, tumor necrosis factor, alpha-1 antitrypsin, anti-thrombin III, leukemia inhibitory factor, glucagon, Protein C, protein kinase C, stem cell factor, follicle stimulating hormone β, urokinase, a nerve growth factor, an insulin-like

growth factor, insulinotropin, parathyroid hormone, lactoferrin, a complement inhibitor, platelet derived growth factor, keratinocyte growth factor, hepatocyte growth factor, endothelial cell growth factor, neurotropin-3, thrombopoietin, chorionic gonadotropin, thrombomodulin, alpha glucosidase, epidermal growth factor, fibroblast growth factor, a cell surface receptor, a transmembrane ion channel, a cholesterol receptor, a receptor for a lipoprotein, an integrin, a cytoskeletal anchoring protein, an immunoglobulin receptor, and a CD antigen.

95. A gene expression product produced by the method of claim 87 or claim 89, wherein said gene expression product is an erythropoietin protein.

96. A gene expression product produced by the method of claim 87 or claim 89, wherein said gene expression product is a growth hormone protein.

97. A gene expression product produced by the method of claim 87 or claim 89, wherein said gene expression product is a G-CSF protein.

98. A method for over-expressing an endogenous gene in a cell *in vivo,* comprising:

(a) introducing a vector comprising a transcriptional regulatory sequence and one or more amplifiable markers into a cell;

(b) integrating said vector into the genome of said cell by non-homologous recombination;

(c) over-expressing an endogenous gene or a portion thereof in said cell by upregulation of said gene by said transcriptional regulatory sequence;

(d) screening said cell for over-expression of said endogenous gene; and

(e) introducing said isolated and cloned cell into an animal under conditions favoring the overexpression of said endogenous gene by said cell *in vivo*.

99. The method of any one of claims 83, 85-87, 89 and 98, wherein said transcriptional regulatory sequence is a promoter.

100. The method of claim 99, wherein said promoter is a viral promoter.

101. The method of claim 100, wherein said viral promoter is the cytomegalovirus immediate early promoter.

102. The method of claim 99, wherein said promoter is a non- viral promoter.

103. The method of claim 99, wherein said promoter is inducible.

104. The method of any one of claims 83, 85-87, 89 and 98, further comprising introducing double strand breaks into the genomic DNA of said cell prior to or simultaneously with integration of said vector.

105. The method of claim 76, further comprising introducing double strand breaks into the genomic DNA of said cell prior to or simultaneously with integration of said vector.

106. The method of claim 77, further comprising introducing double strand breaks into the genomic DNA of said cell prior to or simultaneously with integration of said vector.

107. A gene expression product produced by the method of any one of claims 83, 85-87, 89 and 98.

108. The method of any one of claims 83, 85-87, 89 and 98, wherein said vector construct is linear.

109. A method for producing an expression product of an endogenous gene in a cell comprising:

(a) introducing a vector comprising a transcriptional regulatory sequence into at least one isolated genome-containing cell;

(b) integrating said vector into the genome of said cell by non-homologous recombination;

(c) over-expressing an endogenous gene or a portion thereof in said cell by upregulation of said gene by said transcriptional regulatory sequence;

(d) screening said cell for over-expression of said endogenous gene; and

(e) culturing said cell in reduced serum medium.

110. A method of protein discovery comprising:

(a) introducing a vector comprising a transcriptional regulatory sequence into at least one isolated genome-containing cell;

(b) integrating said vector into the genome of said cell by non-homologous recombination;

(c) culturing said cell in reduced serum medium under conditions that allow over-expression of an endogenous gene or a portion thereof in said cell by upregulation of said gene by said transcriptional regulatory sequence, thereby producing cell-conditioned media; and

(d) screening said cell-conditioned media for the presence of the expression product of said gene or portion thereof.

111. The method of claim 110, further comprising concentrating said cell-conditioned media prior to screening in (d).

112. The method of any one of claims 109-111, wherein said method comprises a high-throughput assay.

113. A method for producing an expression product of an endogenous cellular gene comprising:

(a) introducing a vector comprising a transcriptional regulatory sequence into a cell;

(b) integrating said vector into the genome of said cell by non-homologous recombination;

(c) over-expressing an endogenous gene or a portion thereof in said cell by upregulation of said gene by said transcriptional regulatory sequence;

(d) screening said cell for over-expression of said endogenous gene; and

(e) culturing said cell under conditions favoring the production of the expression product of said endogenous gene by said cell; and

(f) isolating said expression product from a cell mass equivalent to at least 10 liters of cells at 10^4 cells/ml.

114. The method of any of claims 109-111 and 113, wherein said vector further comprises one or more amplifiable markers.

115. The method of any of claims 109-111 and 113, wherein said vector further comprises an unpaired splice donor site.

116. A method for increasing expression of an endogenous gene in a cell *in situ*, the phenotype of said gene being known, without making use of any sequence information of the gene, the method comprising the steps of:

(a) constructing a vector comprising an amplifiable marker, a transcriptional regulatory sequence, and an unpaired splice donor sequence;

(b) delivering copies of the vector to a plurality of cells;

(c) culturing the cells under conditions permitting nonhomologous recombination events between the inserted vector and the genome of the cells;

(d) screening the recombinant cells by assay for the phenotype of said endogenous gene to identify cells in which the expression of said gene has been enhanced; and

(e) selecting for cells with increased expression of said amplifiable marker and said endogenous gene.

117. The method of claim 116, wherein the phenotype is production of a particular protein and the assay is conducted by testing for increased production of the protein.

118. An isolated cell comprising in its genome an inserted genetic construct, said genetic construct comprising an amplifiable marker and a transcriptional regulatory sequence, wherein said construct is inserted into a gene or an upstream region of a gene and activates the expression of said gene, and wherein said gene and upstream region of said gene have no nucleotide sequence homology to said genetic construct.

119. The isolated cell of claim 118, wherein said genetic construct further comprises an exon-unpaired splice donor sequence.

120. The isolated cell of claim 118 or claim 119, wherein said gene encodes a protein selected from the group of proteins consisting of erythropoietin, insulin, growth hormone, glucocerebrosidase, tissue plasminogen activator, granulocyte-colony stimulating factor (G-CSF), granulocyte/macrophage colony stimulating factor (GM-CSF), macrophage colony-stimulating factor (M-CSF), interferon α, interferon β, interferon γ, interleukin-2, interleukin-3, interleukin-4,

interleukin-6, interleukin-8, interleukin-10, interleukin-11, interleukin-12, interleukin-13, interleukin-14, TGF-β, blood clotting factor V, blood clotting factor VII, blood clotting factor VIII, blood clotting factor IX, blood clotting factor X, TSH-β, bone growth factor-2, bone growth factor-7, tumor necrosis factor, alpha-1 antitrypsin, anti-thrombin III, leukemia inhibitory factor, glucagon, Protein C, protein kinase C, stem cell factor, follicle stimulating hormone β, urokinase, a nerve growth factor, an insulin-like growth factor, insulinotropin, parathyroid hormone, lactoferrin, a complement inhibitor, platelet derived growth factor, keratinocyte growth factor, hepatocyte growth factor, endothelial cell growth factor, neurotropin-3, thrombopoietin, chorionic gonadotropin, thrombomodulin, alpha glucosidase, epidermal growth factor, fibroblast growth factor, a cell surface receptor, a transmembrane ion channel, a cholesterol receptor, a receptor for a lipoprotein, an integrin, a cytoskeletal anchoring protein, an immunoglobulin receptor, and a CD antigen.

121. The isolated cell of claim 118 or claim 119, wherein said gene encodes an erythropoietin protein.

122. The isolated cell of claim 118 or claim 119, wherein said gene encodes a growth hormone protein.

123. The isolated cell of claim 118 or claim 119, wherein said gene encodes a G-CSF protein.

124. A method for enhancing expression of a gene, comprising:

(a) introducing a vector into the genome of a cell, said vector containing an enhancer sequence and one or more amplifiable markers, wherein said vector contains no gene-specific targeting sequences;

(b) screening said cell for expression of an endogenous gene; and

(c) selecting for cells with increased expression of said amplifiable marker and said endogenous gene.

125. The method of claim 124, further comprising isolating the cell in which expression of said endogenous gene has been increased.

126. A method for enhancing expression of an endogenous gene in a cell, comprising:

(a) integrating a vector into a cell by non-homologous recombination, said vector comprising an enhancer sequence and one or more amplifiable markers;

(b) screening for nonhomologous recombinant cells that express said endogenous gene, wherein said gene and the upstream and downstream regions of said gene, in which regions said enhancer sequence is active, are not homologous to said vector; and

(c) selecting for cells with increased expression of said amplifiable marker and said endogenous gene.

127. An isolated cell comprising in its genome an inserted artificial genetic construct, the genetic construct comprising one or more amplifiable markers and an enhancer effective at enhancing the expression of a gene in said cell, wherein said genetic construct is inserted into a gene or upstream or downstream regions of a gene, and wherein said gene and said regions, in which said enhancer sequence is active, are not homologous to said genetic construct.

128. The method of any one of claims 80, 83, 85-87, 89, 98, 109-111 and 113, wherein said endogenous gene encodes a transmembrane protein.

129. The method of any one of claims 116, 124 and 126, wherein said gene encodes a cellular transmembrane protein.

130. The method of any one of claims 85, 86, 89 and 98, further comprising isolating and cloning said cell prior to introducing said cell into an animal.

131. The method of any one of claims 85, 86, 89 and 98, wherein said animal is a mammal.

132. The method of claim 131, wherein said mammal is a human.

133. A method for identifying a cell expressing an endogenous gene encoding an integral membrane protein, comprising:

(a) introducing into a cell a vector comprising:

(i) a transcriptional regulatory sequence operably linked to an exonic sequence containing a start codon,

(ii) a signal sequence, and

(iii) an epitope tag followed by an unpaired splice donor site;

(b) integrating said vector into the genome of said cell by non-homologous recombination;

(c) over-expressing an endogenous gene or a portion thereof in said cell by upregulation of said gene by said transcriptional regulatory sequence; and

(d) screening said cell for expression of said epitope tag on the surface of said cell.

134. A method for identifying a cell expressing an endogenous gene encoding an integral membrane protein, comprising:

(a) isolating genomic DNA from a eukaryotic host cell;

(b) combining said isolated genomic DNA with a vector to form a genomic DNA-vector complex, said vector comprising:

(i) a transcriptional regulatory sequence operably linked to an exonic sequence containing a start codon,

(ii) a signal sequence, and

(iii) an epitope tag;

(c) introducing said genomic DNA-vector complex into a eukaryotic host cell;

(d) over-expressing an endogenous gene in said cell by upregulation of said gene by said transcriptional regulatory sequence; and

(e) screening said cell for expression of said epitope tag on the surface of said cell.

135. A method for identifying a cell expressing an endogenous gene encoding an integral membrane protein, comprising:

(a) preparing cDNA from a eukaryotic host cell;

(b) combining said isolated cDNA with a vector to form a cDNA-vector complex, said vector comprising:

(i) a transcriptional regulatory sequence operably linked to an exonic sequence containing a start codon,

(ii) a signal sequence, and

(iii) an epitope tag followed by an unpaired splice donor site;

(c) introducing said cDNA-vector complex into a eukaryotic host cell;

(d) over-expressing an endogenous gene in said cell by upregulation of said gene by said transcriptional regulatory sequence; and

(e) screening said cell for expression of said epitope tag on the surface of said cell.

136. The method of any one of claims 133-135, further comprising isolating said cell expressing said epitope tag.

137. The method of claim 136, further comprising isolating said over-expressed endogenous gene from said isolated cell.

138. A vector comprising:

(a) a first promoter operably linked to an exon and an unpaired splice donor site, and

(b) a second promoter operably linked to a selectable marker lacking a polyadenylation signal.

139. The vector of claim 138, wherein said first and second promoters are present in said vector in the same orientation.

140. The vector of claim 139, wherein said vector is linear and wherein said selectable marker is located 3' to said first promoter.

141. The vector of claim 139, wherein said vector is linear and wherein said second promoter is located 5' to said unpaired splice donor site.

142. The vector of claim 138, wherein said exon lacks a translation start codon.

143. The vector of claim 138, wherein said exon comprises a translation start codon.

144. The vector of claim 138, wherein said exon comprises a translation start codon and a signal secretion sequence.

145. A vector construct comprising:

(a) a first promoter;

(b) a positive selectable marker;

(c) a negative selectable marker; and

(d) an unpaired splice donor site,

wherein said positive and negative selectable markers and said splice donor site are oriented in said vector construct in an orientation that, when said vector construct is integrated into the genome of a eukaryotic host cell in such a way that splicing occurs between said vector-encoded splice donor site and a genome-encoded splice acceptor site, then said positive selectable marker is expressed in active form and said negative selectable marker is either not expressed or is expressed in inactive form.

146. The vector of claim 145, wherein said positive and negative selectable markers are present as a fusion gene.

147. The vector of claim 145, wherein said positive selectable marker, said negative selectable marker, or both said positive and negative selectable markers, lacks a polyadenylation site.

148. The vector of claim 145, wherein said vector further comprises a second promoter operably linked to a second unpaired splice donor site.

149. A vector comprising a first promoter and a second promoter, said first and second promoters being oriented in the same direction, wherein:

(a) said first promoter, but not said second promoter, is operably linked to an unpaired splice donor site; and

(b) said vector comprises no polyadenylation signals downstream of either said first promoter or said second promoter.

150. The vector of claim 149, wherein said vector is linear and wherein said second promoter is located 3' to said first promoter.

151. A vector comprising:

(a) a first promoter operably linked to a first selectable marker containing an unpaired splice donor site; and

(b) a second promoter operably linked to a second selectable marker,

wherein neither said first selectable marker nor said second selectable marker contains a polyadenylation signal.

152. The vector of claim 151, wherein said first and second selectable markers are positive selectable markers.

153. The vector of claim 151, wherein said first selectable marker is located upstream of said second selectable marker.

154. A vector comprising:

(a) a first promoter operably linked to a first exon and a first unpaired splice donor site; and

(b) a second promoter operably linked to a second exon and a second unpaired splice donor site,

wherein the nucleotide sequence of said first exon is different from the nucleotide sequence of said second exon.

155. The vector of claim 154, wherein said first and second exons each comprises a translation start codon and an open reading frame that does not terminate with a stop codon.

156. The vector of claim 154, wherein said first exon, said second exon, or both said first and second exons, lack a translation start codon.

157. A vector construct comprising:

(a) a first promoter operably linked to a positive selectable marker;

(b) a second promoter operably linked to a negative selectable marker; and

(c) an unpaired splice donor site,

wherein said positive and negative selectable markers and said splice donor site are oriented in said vector construct in an orientation that, when said vector construct is integrated into the genome of a eukaryotic host cell in such a way that an endogenous gene in said genome is transcriptionally activated, then said positive selectable marker is expressed in active form and said negative selectable marker is either not expressed or is expressed in inactive form.

158. The vector construct of claim 157, further comprising a third promoter operably linked to a second unpaired splice donor site.

159. The vector of any one of claims 1, 5-7, 10, 58, 59, 62, 138, 145, 149, 151, 154, 157 and 158, said vector further comprising one or more transposition signals.

160. The vector of any one of claims 138, 145, 149, 151, 154, 157 and 158, said vector further comprising one or more amplifiable markers.

161. The vector of any one of claims 1, 5-7, 10, 58, 59, 62, 138, 145, 149, 151, 154, 157 and 158, said vector further comprising one or more viral origins of replication.

162. The vector of any one of claims 1, 5-7, 10, 58, 59, 62, 138, 145, 149, 151, 154, 157 and 158, said vector further comprising one or more viral replication factor genes.

163. The vector of claim 160, wherein said amplifiable marker is selected from the group consisting of dihydrofolate reductase, adenosine deaminase, aspartate transcarbamylase, dihydro-orotase, and carbamyl phosphate synthase.

164. The vector of claim 161, wherein said viral origin of replication is selected from the group consisting of Epstein Barr virus ori P and SV40 ori.

165. The vector of any one of claims 1, 5-7, 10, 58, 59, 62, 138, 145, 149, 151, 154, 157 and 158, said vector further comprising genomic DNA.

166. A host cell comprising the vector of any one of claims 58, 59, 62, 138, 145, 149, 151, 154, 157 and 158.

167. A host cell comprising the vector of claim 159.

168. A host cell comprising the vector of claim 160.

169. A host cell comprising the vector of claim 161.

170. A host cell comprising the vector of claim 162.

171. A host cell comprising the vector of claim 165.

172. The host cell of claim 166, wherein said host cell is an isolated cell.

173. The host cell of any one of claims 167-171, wherein said host cell is an isolated cell.

174. A library of cells comprising the vector of any one of claims 1, 5-7, 10, 58, 59, 62, 138, 145, 149, 151, 154, 157 and 158.

175. A library of cells comprising the vector of claim 159.

176. A library of cells comprising the vector of claim 160.

177. A library of cells comprising the vector of claim 161.

178. A library of cells comprising the vector of claim 162.

179. A library of cells comprising the vector of claim 165.

180. A method for activation of an endogenous gene in a cell comprising:

(a) transfecting a genome-containing cell with the vector of any one of claims 1, 5-7, 10, 58, 59, 62, 138, 145, 149, 151, 154, 157 and 158; and

(b) culturing said cell under conditions suitable for nonhomologous integration of said vector into the genome of said cell, wherein said integration results in the activation of an endogenous gene in the genome of said cell.

181. A method for identifying a gene comprising:

(a) transfecting a plurality of genome-containing cells with the vector of any one of claims 1, 5-7, 10, 58, 59, 62, 138, 145, 149, 151, 154, 157 and 158;

(b) culturing said cells under conditions suitable for nonhomologous integration of the vector into the genome of the host cell;

(c) selecting for cells in which said vector has integrated into the genomes of said cells;

(d) isolating RNA from said selected cells;

(e) producing cDNA from said isolated RNA; and

(f) identifying a gene in said cDNA by isolating one or more cDNA molecules containing one or more nucleotide sequences from said vector.

182. The method of claim 181, wherein said identification in (f) is accomplished by hybridizing said cDNA to said vector.

183. The method of claim 181, wherein said identification in (f) is accomplished by sequencing said cDNA and comparing the nucleotide sequence of said cDNA to the nucleotide sequence of said vector.

184. The vector of claim 151, wherein said unpaired splice donor site is positioned upstream of, or within, said first selectable marker such that, when said vector is integrated into the genome of a eukaryotic host cell resulting in splicing from said unpaired splice donor site to a genome-encoded splice acceptor site, then said first selectable marker is expressed in inactive form or is not expressed at all.

185. A method for isolating cells in which a single exon gene has been activated, comprising:

(a) transfecting a plurality of genome-containing eukaryotic cells with the vector of claim 184;

(b) culturing said cells under conditions suitable for nonhomologous integration of the vector into the genomes of said cells; and

(c) selecting for cells in which said first and second selectable markers are expressed in their active forms.

186. The method of claim 185, further comprising:

(d) isolating RNA from the selected cells;

(e) producing cDNA from said isolated RNA; and

(f) isolating a single exon gene from said cDNA.

187. A method for isolating exon I of a gene comprising:

(a) transfecting one or more genome-containing eukaryotic cells with the vector of any one of claims 138, 139, 141, 149, 151 and 154;

(b) culturing said cells under conditions suitable for nonhomologous integration of the vector into the genome of said cells;

(c) selecting for cells in which said vector has transcriptionally activated an endogenous gene containing one or more exons;

(d) isolating RNA from said selected cells;

(e) producing cDNA from said isolated RNA;

(f) recovering cDNA molecules containing a first exon from said vector spliced to a second exon from said endogenous gene, thereby obtaining one or more vector exon-tagged cDNA molecules; and

(g) using said vector exon-tagged cDNA molecules to recover the activated endogenous gene containing exon I.

188. A method for expressing a transcript containing exon I of a gene, said method comprising:

(a) transfecting one or more genome-containing eukaryotic cells with the vector of any one of claims 138, 139, 141, 149, 151 and 154;

(b) culturing said cells under conditions suitable for nonhomologous integration of the vector into the genome of said cells; and

(c) culturing said cells under conditions suitable for expression of a transcript containing exon I from an endogenous gene.

189. A method for producing a gene product comprising:

(a) isolating genomic DNA, containing at least one gene, from a eukaryotic cell;

(b) inserting into said isolated genomic DNA, by *in vitro* transposition, a vector comprising one or more transposition signals, one or more promoters, one or more exons, and one or more unpaired splice donor sites, thereby forming a genomic DNA-vector complex;

(c) introducing said genomic DNA-vector complex into a eukaryotic host cell; and

(d) culturing said host cell under conditions suitable for expression of said gene.

190. The method of claim 189, further comprising isolating an expression product of said gene.

191. A method for producing a gene product encoded by an endogenous cellular genomic gene, comprising:

(a) isolating genomic DNA, containing at least one gene, from a eukaryotic cell;

(b) inserting into or otherwise combining with said isolated genomic DNA, the vector of any one of claims 138, 139, 141, 149, 151 and 154, thereby producing a vector-genomic DNA complex;

(c) transfecting said vector-genomic DNA complex into a suitable eukaryotic host cell; and

(d) culturing said host cell under conditions suitable to result in transcription of one or more genes encoded by said vector contained in said vector-genomic DNA complex.

192. The method of claim 191, further comprising:

(e) isolating RNA produced by said transcription from said host cell;

(f) producing one or more cDNA molecules from said isolated RNA; and

(g) recovering one or more cDNA molecules containing vector sequences at the 5' ends of said cDNA molecules, thereby isolating said gene.

193. The method of claim 191, wherein said vector further comprises one or more transposition signals, and wherein said vector is inserted into said isolated genomic DNA by *in vitro* transposition.

194. The method of claim 191, wherein said isolated genomic DNA is present in a cloning vector.

195. A method for producing a protein comprising:

(a) isolating genomic DNA from one or more cells;

(b) inserting into or otherwise combining with said isolated genomic DNA, the vector of any one of claims 138, 139, 141, 149, 151 and 154, thereby producing a vector-genomic DNA complex;

(c) transfecting said vector-genomic DNA complex into a suitable host cell; and

(d) culturing said cell under conditions suitable to result in protein expression from said genomic DNA contained in said vector-genomic DNA complex.

196. A method for producing a protein comprising:

(a) isolating genomic DNA from one or more cells;

(b) integrating a vector comprising one or more transposition signals and a transcriptional regulatory sequence operably linked to an exon-unpaired splice donor complex, into said isolated genomic DNA by transposition, thereby producing a vector-genomic DNA complex;

(c) transfecting said vector-genomic DNA complex into a suitable host cell; and

(d) culturing said cell under conditions suitable to result in protein expression from said genomic DNA contained in said vector-genomic DNA complex.

197. A method for expressing a gene, comprising:

(a) isolating genomic DNA, containing one or more genes, from one or more eukaryotic cells;

(b) combining said isolated genomic DNA with a vector comprising:

(i) a selectable marker,

(ii) a transcriptional regulatory sequence operably linked to a translational start codon,

(iii) a secretion signal sequence,

(iv) an epitope tag, and

(v) an unpaired splice donor site,

thereby producing a vector-genomic DNA complex;

(c) introducing said vector-genomic DNA complex into a cell;

(d) selecting for cells containing said vector-genomic DNA complex; and

(e) culturing said cell under conditions suitable to result in expression of a gene contained in said vector-genomic DNA complex.

198. The method of claim 195, wherein said host cell is selected for a cell containing said transfected vector-genomic DNA complex prior to, during, or following being cultured under conditions suitable to result in protein expression.

199. The method of claim 196, wherein said vector further comprises a selectable marker, and wherein said host cell is selected for a cell containing said transfected vector-genomic DNA complex prior to being cultured under conditions suitable to result in protein or gene expression.

200. The method of claim 194, wherein said cloning vector is selected from the group consisting of a BAC, a YAC, a PAC, a cosmid, a phage, and a plasmid.

201. The method of claim 191, further comprising isolating said protein.

202. A protein produced by the method of claim 195.

203. A protein produced by the method of any one of claims 197-199.

204. A protein produced by the method of claim 201.

205. A method for protein expression comprising:

(a) transfecting a host cell with a vector comprising a heterologous promoter operably linked to:

(i) a heterologous exon,

(ii) a heterologous splice donor site,

(iii) a genomic DNA fragment encoding a gene or portion thereof, and

(iv) one or more selectable markers,

wherein said heterologous exon either lacks a translation start codon or encodes a translation start codon and an open reading frame that is not terminated by a stop codon;

(b) selecting for a cell containing said transfected vector; and

(c) culturing said selected transfected host cell under conditions suitable for protein expression from said vector.

206. The method of claim 205, wherein said vector further comprises a viral origin of replication.

207. The method of claim 206, wherein said viral origin of replication is Epstein Barr Virus oriP.

208. A vector comprising:

(a) a heterologous promoter;

(b) a heterologous exon;

(c) a heterologous splice donor site;

(d) a genomic fragment encoding a gene or portion thereof;

(e) one or more selectable markers; and

(f) one or more viral origins of replication,

wherein said heterologous exon either lacks a translation start codon or encodes a translation start codon and an open reading frame that is not terminated by a stop codon, and wherein said genomic fragment is oriented downstream of said heterologous promoter, said exon and said splice donor site, such that upon introduction of said vector into a host cell, protein is expressed from said gene or portion thereof encoded by said genomic fragment.

209. The vector of claim 208, wherein said selectable marker lacks a polyadenylation signal.

210. The vector of claim 208, further comprising one or more genes encoding one or more viral replication proteins.

211. The vector of claim 208, further comprising an amplifiable marker.

212. A cell comprising the vector of any one of claims 208-211.

213. The cell of claim 212, wherein said cell is an isolated cell.

214. The vector construct of claim 10, wherein said first transcriptional regulatory sequence is in the same orientation in said vector construct as said second transcriptional regulatory sequence.

215. The vector construct of claim 145 or 157, wherein said positive selectable marker is selected from the group consisting of a neomycin gene, a hypoxanthine phosphribosyl transferase gene, a puromycin gene, a dihydrooratase gene, a glutamine synthetase gene, a histidine D gene, a carbamyl phosphate synthase gene, a dihydrofolate reductase gene, a multidrug resistance 1 gene, an aspartate transcarbamylase gene, a xanthine-guanine phosphoribosyl transferase gene, and an adenosine deaminase gene.

216. The vector construct of claim 145 or 157, wherein said negative selectable marker is selected from the group consisting of a hypoxanthine phosphribosyl transferase gene, a thymidine kinase gene, and a diphtheria toxin gene.

217. The vector of claim 157, wherein said negative selectable marker is located upstream of said positive selectable marker.

218. A host cell stably expressing a protein, wherein said host cell comprises a vector comprising a promoter, an exon/splice donor complex, and a genomic fragment encoding said protein or portion thereof, wherein said promoter and exon/splice donor complex are heterologous to said genomic fragment.

219. The host cell of claim 218, wherein said vector is integrated into the genome of said cell.

220. The host cell of claim 218, wherein said vector further comprises a viral origin of replication and is maintained within said host cell as an episome.

221. The cell of claim 217 or claim 219, wherein said vector further comprises one or more selectable markers.

222. The cell of claim 219, wherein said viral origin of replication is Epstein Barr Virus oriP.

223. A method for activating expression from an endogenous gene comprising:

(a) introducing into a chromosome-containing host cell a vector suitable for activating an endogenous gene;

(b) treating said cell with an agent capable of introducing DNA breaks in the chromosome of said host cell prior to or following introduction of said vector; and

(c) integrating said vector into said DNA breaks so as to result in the formation of an operable linkage between said vector and said endogenous gene, whereby said endogenous gene is activated by one or more vector-encoded nucleotide sequences.

224. The method of claim 223, wherein said activation in (d) is accomplished by isolating said host cell and culturing said host cell under conditions favoring activation of said endogenous gene.

225. A vector comprising:

(a) a transcriptional regulatory sequence operably linked to a gene;

(b) a viral origin of replication; and

(c) an amplifiable marker.

226. A method for increasing expression of a gene, comprising:

(a) introducing the vector of claim 225 into a host cell, wherein said vector is maintained as an episome within said host cell; and

(b) selecting for increased expression of said amplifiable marker and said gene.

227. A method for cleaving a cDNA molecule derived from an unspliced cellular transcript molecule, comprising:

(a) integrating the vector of claim 5 or claim 7 into the genome of one or more eukaryotic host cells;

(b) culturing said host cell under conditions suitable for expression from said transcriptional regulatory sequence;

(c) isolating RNA from said host cell;

(d) producing cDNA from said isolated RNA; and

(e) digesting said cDNA with an enzyme that cleaves at said rare cutting restriction site.

228. A method for drug discovery comprising:

(a) integrating a vector into the genome of a eukaryotic host cell, wherein said vector integration activates expression of an endogenous gene in said host cell;

(b) culturing said cell under conditions favoring expression of said activated gene, thereby producing a gene product of said activated gene;

(c) treating said cell with one or more test compounds to be screened for drug activity; and

(d) determining the ability of said one or more test compounds to interact with, or affect a cellular phenotype induced by, said gene product.

229. A method for drug discovery comprising:

(a) integrating a vector into the genome of a eukaryotic host cell, wherein said vector integration activates expression of an endogenous gene in said host cell;

(b) culturing said cell in reduced serum medium under conditions favoring production of a gene product of said activated gene, thereby producing cell-conditioned media comprising said gene product; and

(c) screening one or more test compounds for drug activity by determining the ability of said test compounds to interact with said gene product in said cell-conditioned media.

230. The method of claim 229, further comprising concentrating said cell-conditioned media prior to said screening in (c).